

Seminar Hasil Tesis

PEMODELAN REGRESI TOBIT KUANTIL BAYESIAN PADA PENGELUARAN RUMAH TANGGA UNTUK KONSUMSI SUSU

Oleh:

Evellin Dewi Lusiana
(1313201010)

Dosen Pembimbing: Dr. Dra. Ismaini Zain, M.Si

Latar Belakang [1]

Pemodelan regresi →
variabel respon
memiliki proporsi
observasi bernilai nol
cukup besar

-Regresi
Logistik/Probit
- Regresi Linier

Sebagian
informasi hilang

(Greene, 2001)

Regresi Tobit

Semua observasi
diikutsertakan

Regresi Tobit →
Metode MLE

Conditional Mean

(Greene, 2001)

Uzunoz dan Akcay
(2012), Babolian dan AB
Karim (2010)

Komrattanapanya
(2013), Aisyah et al.
(2012), Zain et. al
(2000), Zain (1997)

Latar Belakang [2]

- Susenas : konsumsi susu
- Faktor-faktor sosio-ekonomi dan demografi (Wham & Worsley, 2003)

Aplikasi

?(3)

kajian

?(1)

Pengeluaran RT untuk kesehatan di AS

(Yue dan Hong, 2012)

aplikasi

Regresi kuantil tersensor

(Powell, 1986)

Metode BRCENS (Fitzenberger dan Winker, 2007)

(Greene, 2001)

?(2)

perbandingan

Data tersensor

Regresi Tobit → Metode MLE

Conditional Mean
Conditional Quantile

Regresi tobit kuantil bayesian

(Yu dan Stander, 2007)

MCMC: Metropolis-Hastings

Rumusan Masalah

Bagaimana mengkaji estimator parameter dan pemilihan model regresi tobit kuantil terbaik dengan pendekatan bayesian?

Bagaimana perbandingan performa estimator parameter model tobit kuantil bayesian bila dibandingkan dengan estimator Powell?

Bagaimana model regresi tobit kuantil bayesian terbaik bagi pengeluaran rumah tangga untuk konsumsi susu?

Tujuan

Mengkaji estimasi parameter dan pemilihan model regresi tobit kuantil terbaik dengan pendekatan bayesian

Membandingkan performa estimator parameter model tobit kuantil bayesian bila dibandingkan dengan estimator Powell menggunakan teknik simulasi

Memodelkan regresi tobit kuantil bayesian terbaik bagi pengeluaran rumah tangga untuk konsumsi susu

Manfaat Penelitian

- Mengembangkan wawasan dan pengetahuan mengenai analisis **regresi kuantil** pada umumnya dan **regresi tobit kuantil bayesian** pada khususnya serta **pemilihan model terbaiknya**
- Memperoleh informasi mengenai seberapa baik **performa** dari **model regresi kuantil bayesian** bila digunakan dalam berbagai kondisi data yang tersensor dan dibandingkan dengan **estimator Powell**
- Memberikan informasi yang lebih lengkap tentang **model pengeluaran rumah tangga untuk konsumsi susu**, sehingga diharapkan bisa membantu upaya pengambil kebijakan untuk meningkatkan konsumsi susu di suatu wilayah.

Batasan Masalah

- Teknik MCMC yang digunakan adalah algoritma **Metropolis-Hastings**.
- Ukuran untuk mengevaluasi performa estimator adalah **RMSE**.
- Estimator **Powell** yang akan dibandingkan dengan estimator parameter model regresi **tobit kuantil bayesian** diperoleh menggunakan metode BRCENS.
- Kuantil yang akan dimodelkan adalah **0.05, 0.25, 0.50, 0.75** dan **0.95**.
- Pemilihan model terbaik dilakukan berdasarkan **Bayes faktor**

Model Regresi Tobit

- Suatu variabel respon Y disebut tersensor pada batas bawah apabila untuk setiap $i=1,2,\dots,n$ berlaku persamaan berikut (McBee (2010); Leiker (2012))

$$y_i = \begin{cases} \tau & y_i^* \leq \tau \\ y_i^* & y_i^* > \tau \end{cases}$$

$$y_i = \begin{cases} \tau & y_i^* \leq \tau \\ \mathbf{x}_i^t \boldsymbol{\beta} + \varepsilon_i & y_i^* > \tau \end{cases}$$

$$y_i^* = f(\mathbf{x}_i) + \varepsilon_i$$
$$f(\mathbf{x}_i) = \mathbf{x}_i^t \boldsymbol{\beta}$$

$$y_i = \text{maks}(y_i^*, \tau)$$

$$y_i = \text{maks}(\mathbf{x}_i^t \boldsymbol{\beta} + \varepsilon_i, \tau)$$

- Estimasi parameter model regresi Tobit menggunakan MLE dan bantuan metode iteratif Newton Raphson (Wooldridge, 2002)

Uji Linieritas

- **Uji Terasvirta** (H_0 : model linier vs H_1 : model nonlinier)

$$F_j = \frac{(SSR_{0j} - SSR_{1j}) / m}{SSR_{1j} / (n - m - 2)} \sim F_{(m, n-m-2)}$$

di mana:

SSR_{0j} = jumlah kuadrat regresi tanpa variabel prediktor ke-j nonlinier tambahan
 SSR_{1j} = jumlah kuadrat regresi dengan variabel prediktor ke-j nonlinier tambahan
 m = banyaknya komponen nonlinier yang ditambahkan dalam model

H_0 ditolak jika $P(F_{(m, n-m-2)} > F_j) < \alpha$ (Subanar dan Suhartono, 2006).

- **Uji Biserial Point** (H_0 : hubungan antar variabel tidak linier vs H_1 : hubungan antar variabel linier)

$$r_{pb} = \frac{\bar{y}_1 - \bar{y}_0}{s_y} \sqrt{\frac{n_1 n_0}{n(n-1)}}$$

$$t = \frac{r_{pb} \sqrt{n-2}}{\sqrt{1-r_{pb}^2}} \sim t_{(n-2)}$$

di mana: \bar{y}_1 = rata-rata respon kategori 1
 \bar{y}_0 = rata-rata respon kategori 0
 n_1 = banyaknya observasi kategori 1
 n_0 = banyaknya observasi kategori 0
 n = $n_1 + n_0$

H_0 ditolak jika $P(t_{(n-2)} > |t|) < \alpha$ (Glass dan Hopkins, 1995).

Uji Beda

- Uji t

di mana:

$$H_0: \mu_1 - \mu_0 = 0$$

$$H_1: \mu_1 - \mu_0 \neq 0$$

$$t = \frac{\bar{x}_1 - \bar{x}_0}{s_{\bar{x}_1 - \bar{x}_0}} \sim t_{(db)}$$

\bar{x}_1 = rata-rata sampel kelompok kategori

\bar{x}_0 = rata-rata sampel kelompok kategori

$s_{\bar{x}_1 - \bar{x}_0}$ = standar error selisih dua rata-rata

H_0 ditolak jika $P(t_{(db)} > |t|) < \alpha$ (Robinson *et al.*, 2013).

- Uji *chi-square*

H_0 : hubungan antar kedua variabel saling bebas

H_1 : hubungan antar kedua variabel tidak saling bebas

$$\chi^2 = \sum_{m=1}^g \frac{(O_m - E_m)^2}{E_m} \sim \chi^2_{(g-1)}$$

di mana:

O_m = frekuensi observasi dalam kategori ke-m

E_m = frekuensi yang diharapkan dalam kategori ke-m

H_0 ditolak jika $P(\chi^2_{(g-1)} > \chi^2) < \alpha$ (Agresti, 2002).
g = banyaknya kategori

Model Regresi Kuantil

- Bentuk umum model regresi kuantil (Buhai, 2005)

$$y_i = \mathbf{x}_i^t \boldsymbol{\beta}(\theta) + \varepsilon(\theta)_i$$

- Estimator model ini merupakan solusi minimasi persamaan berikut, dengan metode pemrograman linier (Koenker dan Machado, 1999)

$$\min_{\boldsymbol{\beta} \in \mathbb{R}^p} \left[\sum_{i \in \{i: y_i \geq \mathbf{x}_i^t \boldsymbol{\beta}\}} \rho_{\theta}(y_i - f(\mathbf{x}_i)) \right]$$

Model Regresi Kuantil Tersensor

- Estimator parameter model regresi kuantil tersensor diperoleh dari minimasi berikut (Powell, 1986).

$$\min_{\beta \in \mathbb{R}^p} \left[\sum_{i=1}^n \rho_{\theta} \left(y_i - maks(\mathbf{x}_i^t \beta(\theta), \tau) \right) \right]$$

- Minimasi tersebut dapat diselesaikan dengan metode pemrograman linier ILPA (Buchinsky, 1994) dan BRCENS (Fitzenberger dan Winker, 2007)

Model Regresi Kuantil Bayesian

- Residual model regresi kuantil diasumsikan berdistribusi asimetris Laplace (Yu dan Stander, 2007)

$$f_{\theta}(\varepsilon_i) = \theta(1-\theta)\exp(-\rho_{\theta}(y_i - \mathbf{x}_i^t\boldsymbol{\beta}(\theta))) \quad , 0 < \theta < 1$$

- Fungsi likelihood dari model ini sebagai berikut.

$$L(y|\boldsymbol{\beta}(\theta)) = \theta^n (1-\theta)^n \exp\left[-\sum_{i=1}^n \rho_{\theta}(y_i - \mathbf{x}_i^t\boldsymbol{\beta}(\theta))\right]$$

- Distribusi posterior $\pi(\boldsymbol{\beta}(\theta)|\mathbf{y})$ dapat diperoleh dengan kaidah Bayes

$$\pi(\boldsymbol{\beta}(\theta)|\mathbf{y}) \propto L(\mathbf{y}|\boldsymbol{\beta}(\theta))p(\boldsymbol{\beta}(\theta))$$

Teknik MCMC: Algoritma Metropolis-Hastings

1. Menentukan nilai awal $\beta^{(1)}$
2. Untuk $i=2,3,\dots,N$ (N =banyaknya loop chain)
 - 1) Tentukan $\beta = \beta^{(i-1)}$
 - 2) Membangkitkan kandidat nilai parameter baru dari distribusi proposal β' (misal Gaussian) $\pi(\beta'|\beta)$
 - 3) Hitung peluang acceptance $\alpha = \min\left(1, \frac{L(y|\beta')\pi(\beta)}{L(y|\beta)\pi(\beta')}$
 - 4) Perbarui $\beta^{(i)} = \beta'$ dengan peluang acceptance sebesar α

Bayes Faktor

- Bayes faktor dapat didefinisikan sebagai berikut (Kass dan Raftery, 1995)

$$B = \frac{m(\mathbf{y} | M_1)}{m(\mathbf{y} | M_2)} = \frac{\int L(\mathbf{y} | M_1, \boldsymbol{\beta}_1) \pi(\boldsymbol{\beta}_1 | M_1) d\boldsymbol{\beta}_1}{\int L(\mathbf{y} | M_2, \boldsymbol{\beta}_2) \pi(\boldsymbol{\beta}_2 | M_2) d\boldsymbol{\beta}_2}$$

- Interpretasi Bayes faktor (Kass dan Raftery, 1995)

2 log B	B	Kekuatan Pembuktian	Model
< 0	< 1	Negatif	M_1
0 - 2	1 - 3	Tidak ada	
2 - 5	3 - 12	Positif	
5 - 10	12 - 150	Kuat	
> 10	> 150	Sangat kuat	

- Marginal likelihood diestimasi dari output algoritma Metropolis-Hastings (Chib & Jeliakoz, 2001)

Root Mean Square Error (RMSE)

- RMSE (*Root Mean Square Error*) merupakan salah satu statistik yang sering digunakan untuk mengevaluasi kebaikan performa model/estimator. Formulasi RMSE sebagai berikut (Chai dan Draxler, 2014).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n e_i^2}$$

Faktor-faktor yang Memengaruhi Pengeluaran Rumah Tangga untuk Konsumsi Susu



Sumber Data

Informasi	Keterangan
Sumber Data	Survei Sosial Ekonomi Rumah Tangga Indonesia (SUSETI)
Tahun	2011
Pelaksana Survei	Surveymeter
Unit Observasi	Rumah Tangga
Jumlah Sampel	5,432 unit

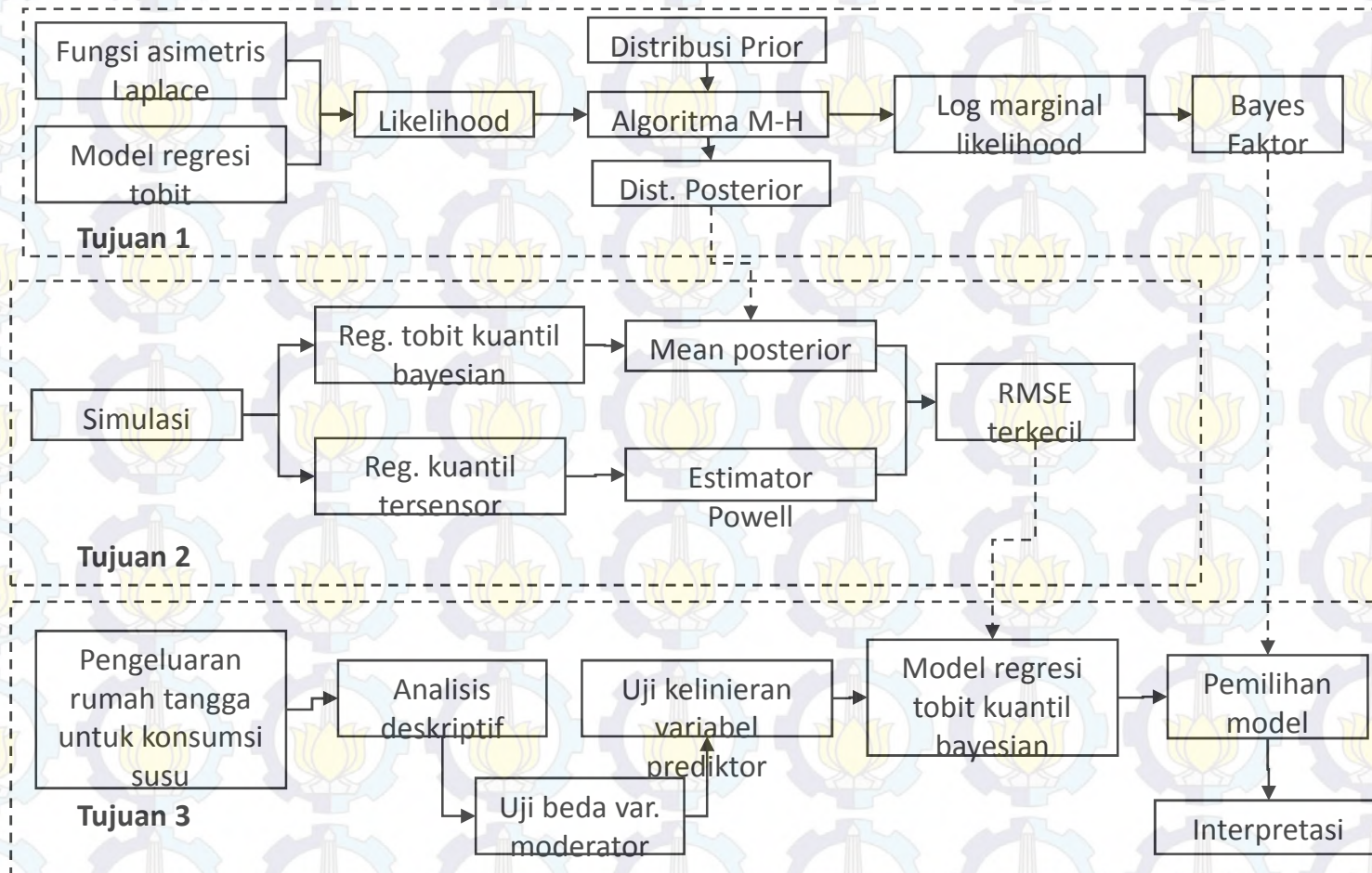
Variabel Penelitian

- Variabel respon → pengeluaran rumah tangga untuk konsumsi susu
- Variabel prediktor
 - X_1 = pendapatan rumah tangga
 - X_2 = tingkat pendidikan kepala rumah tangga (d_{11} , d_{12})
 - X_3 = persentase pengeluaran pangan
 - X_4 = jumlah anggota rumah tangga
 - X_5 = persentase anggota rumah tangga bekerja
 - X_6 = persentase anggota rumah tangga usia ≤ 12 tahun
 - X_7 = rata-rata pengeluaran per kapita
 - Z = wilayah tempat tinggal

Struktur Data

Sampel ke-	Y	X_1	d_{11}	d_{12}	X_3	X_4	X_5	X_6	X_7
1	Y_1	X_{11}	d_{111}	d_{121}	X_{31}	X_{41}	X_{51}	X_{61}	X_{71}
2	Y_2	X_{12}	d_{112}	d_{122}	X_{32}	X_{42}	X_{52}	X_{62}	X_{72}
3	Y_3	X_{13}	d_{113}	d_{123}	X_{33}	X_{43}	X_{53}	X_{63}	X_{73}
4	Y_4	X_{14}	d_{114}	d_{124}	X_{34}	X_{44}	X_{54}	X_{64}	X_{74}
.
.
.
n_z	Y_n	X_{1n}	d_{11n}	d_{12n}	X_{3n}	X_{4n}	X_{5n}	X_{6n}	X_{7n}

Kerangka Penelitian





HASIL DAN PEMBAHASAN

Tujuan 1

- Estimasi Parameter Model TKB

$$y_i = maks(\mathbf{x}_i^t \boldsymbol{\beta} + \varepsilon_i, \tau)$$

$$f_{\theta}(u) = \theta(1 - \theta) \exp(-\rho_{\theta}(u)) \quad , 0 < \theta < 1$$

$$f_{\theta}(\varepsilon_i) = \theta(1 - \theta) \exp\{-\rho_{\theta}(y_i - maks(\mathbf{x}_i^t \boldsymbol{\beta}, 0))\}$$

$$\begin{aligned} L(\mathbf{y} | \boldsymbol{\beta}(\theta)) &= \prod_{i=1}^n f_{\theta}(\varepsilon_i) \\ &= \prod_{i=1}^n \theta(1 - \theta) \exp\{-\rho_{\theta}(y_i - maks(\mathbf{x}_i^t \boldsymbol{\beta}, 0))\} \\ &= \theta^n (1 - \theta)^n \exp\left\{-\rho_{\theta} \sum_{i=1}^n (y_i - maks(\mathbf{x}_i^t \boldsymbol{\beta}, 0))\right\} \end{aligned}$$

$$\pi(\boldsymbol{\beta}(\theta) | \mathbf{y}) \propto L(\mathbf{y} | \boldsymbol{\beta}(\theta)) \pi(\boldsymbol{\beta}(\theta))$$

Tujuan 1: Pemilihan regresi TKB terbaik dengan pendekatan bayesian

- Bayes Faktor TKB

$$\log \hat{m}(\mathbf{y} | M_k) = \log L(\mathbf{y} | M_k, \boldsymbol{\beta}_k^*) + \log \pi(\boldsymbol{\beta}_k^* | M_k) - \log \pi(\boldsymbol{\beta}_k^* | \mathbf{y}, M_k)$$

$$\begin{aligned} \log \hat{m}(\mathbf{y} | M_k) = & n \log \{\theta(1 - \theta)\} - \left\{ \rho_\theta \sum_{i=1}^n (y_i - \max(\mathbf{x}_i^t \boldsymbol{\beta}_k^*, 0)) \right\} \\ & + \log \pi(\boldsymbol{\beta}_k^* | M_k) - \log \pi(\boldsymbol{\beta}_k^* | \mathbf{y}, M_k) \end{aligned}$$

$$\pi(\boldsymbol{\beta}_k^* | \mathbf{y}) = \frac{G^{-1} \sum_{i=1}^G \alpha(\boldsymbol{\beta}^{(g)}, \boldsymbol{\beta}^*) \pi(\boldsymbol{\beta}^{(g)}, \boldsymbol{\beta}^*)}{J^{-1} \sum_{i=1}^J \alpha(\boldsymbol{\beta}^*, \boldsymbol{\beta}^{(j)})}$$

di mana:

$$\alpha(\boldsymbol{\beta}, \boldsymbol{\beta}^*) = \min \left\{ 1, \frac{\pi(\boldsymbol{\beta}^*)}{\pi(\boldsymbol{\beta})} L^*(\boldsymbol{\beta}^*, \boldsymbol{\beta}) \right\}$$

$\{\boldsymbol{\beta}^{(j)}\}$ = sampel yang berasal dari distribusi proposal

$\{\boldsymbol{\beta}^{(g)}\}$ = sampel yang berasal dari distribusi posterior

$$L^*(\boldsymbol{\beta}^*, \boldsymbol{\beta}) = \frac{L(\boldsymbol{\beta}^*)}{L(\boldsymbol{\beta})}$$

$$\begin{aligned} &= \frac{\theta^n (1 - \theta)^n \exp \left\{ -\rho_\theta \sum_{i=1}^n (y_i - \max(\mathbf{x}_i^t \boldsymbol{\beta}^*, 0)) \right\}}{\theta^n (1 - \theta)^n \exp \left\{ -\rho_\theta \sum_{i=1}^n (y_i - \max(\mathbf{x}_i^t \boldsymbol{\beta}, 0)) \right\}} \end{aligned}$$

$$B_{12} = \exp \left\{ \log \hat{m}(\mathbf{y} | M_1) - \log \hat{m}(\mathbf{y} | M_2) \right\}$$

Perbandingan Performa Estimator TKB dan Powell

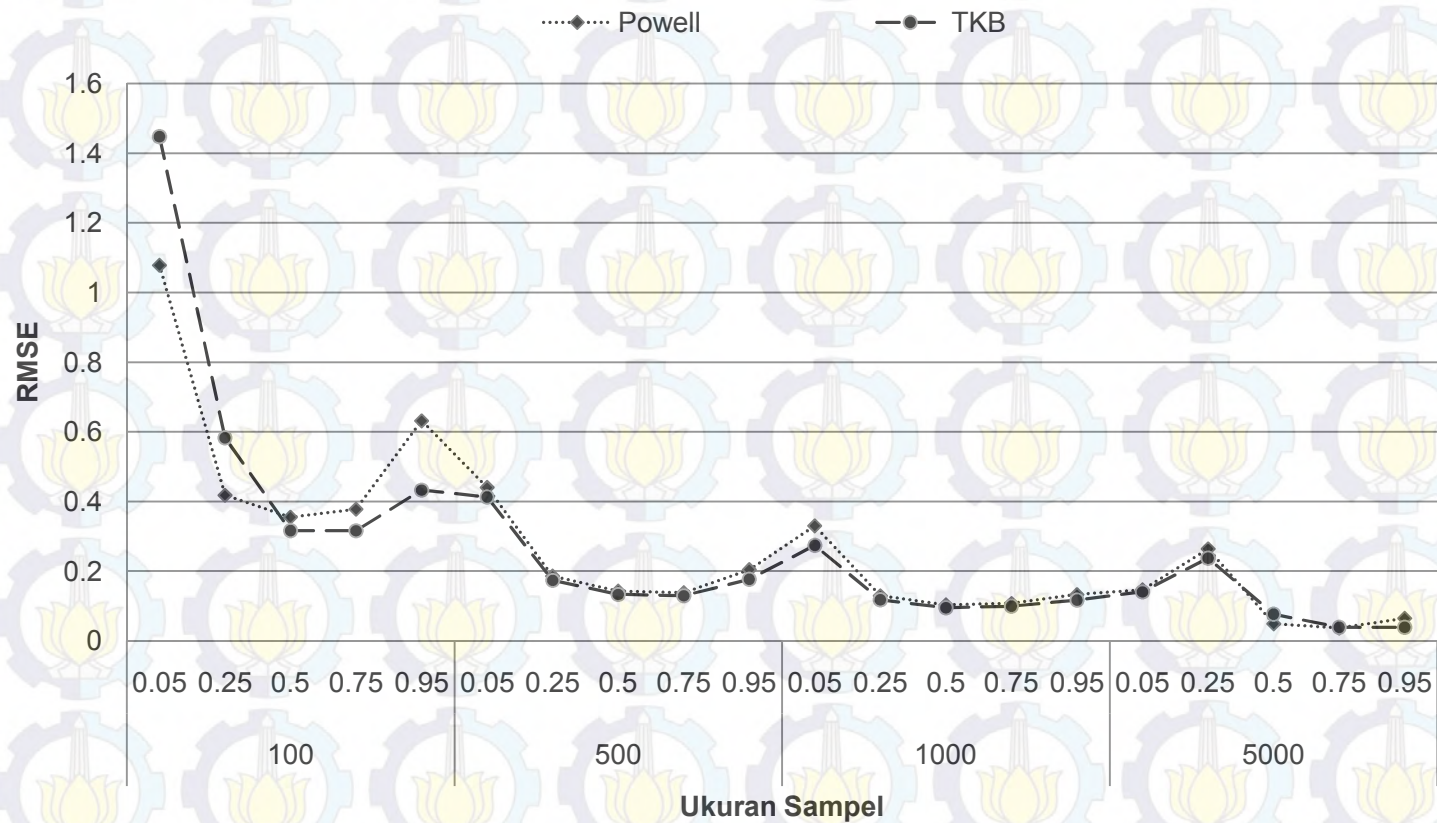
- Jika banyaknya prediktor berbeda

Banyaknya Prediktor (p)	TKB					Powell				
	0.05	0.25	0.50	0.75	0.95	0.05	0.25	0.50	0.75	0.95
1	1.857	1.008	0.274	0.154	0.195	1.288	0.927	0.208	0.151	0.215
3	1.448	0.583	0.316	0.316	0.433	1.079	0.418	0.355	0.378	0.632
5	2.931	0.560	0.480	0.429	0.597	0.869	0.530	0.404	0.440	0.998
7	1.362	0.629	0.458	0.416	0.701	1.264	0.652	0.500	0.541	1.766

Perbandingan Performa Estimator TKB dan Powell

Tujuan 2

- Jika ukuran sampel berbeda



Perbandingan Performa Estimator TKB dan Powell

- Jika distribusi error berbeda

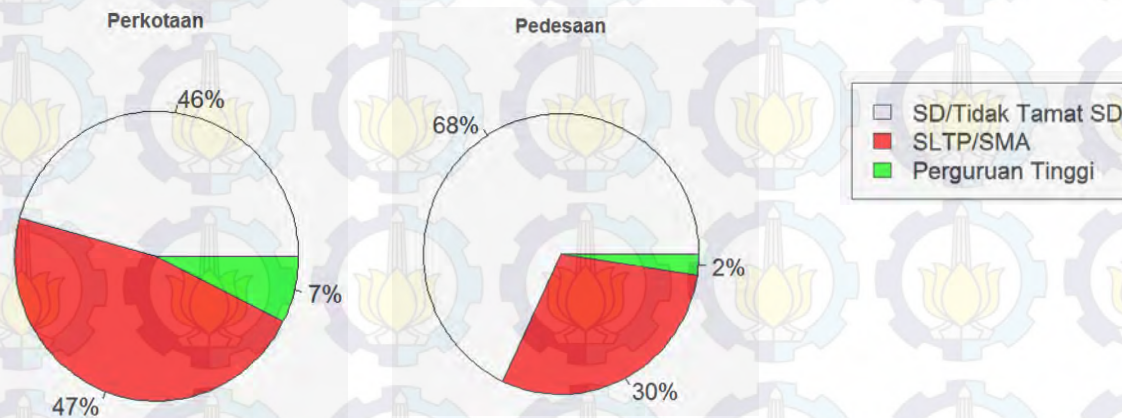
Estimator	Normal			Heteroscedastics			Mixture		
	TKB	Powell	Tobit	TKB	Powell	Tobit	TKB	Powell	Tobit
β_0	0.094	0.105	0.070	0.115	0.125	0.181	0.354	0.355	0.311
β_1	0.050	0.053	0.042	0.064	0.067	0.059	0.055	0.060	0.133
β_2	0.146	0.154	0.113	0.220	0.235	0.311	0.229	0.237	0.349
β_3	0.066	0.065	0.045	0.084	0.087	0.066	0.075	0.077	0.222



Tujuan 3: Pemodelan Regresi TKB bagi Pengeluaran RT untuk konsumsi susu

Deskripsi Data Penelitian

Statistik	Variabel Wilayah	X ₁	X ₃	X ₄	X ₅	X ₆	X ₇	Y
Minimum	Pedesaan	82.00	1.12	1.00	10.00	0.00	10.42	0.00
	Perkotaan	100.00	0.69	2.00	8.33	0.00	15.74	0.00
Maksimum	Pedesaan	80000.0	92.66	18.00	100.00	75.00	4296.15	458.000
	Perkotaan	34000.0	91.52	19.00	100.00	75.00	8416.03	750.000
Rata-rata	Pedesaan	1076.81	50.62	4.85	41.72	25.56	170.26	8.868
	Perkotaan	1780.51	44.65	5.45	37.80	25.48	233.44	19.341
Standar Deviasi	Pedesaan	1888.94	19.73	1.61	17.52	15.81	286.26	23.378
	Perkotaan	2184.92	19.02	1.95	16.49	15.95	439.93	39.786



Tujuan 3: Pemodelan Regresi TKB bagi Pengeluaran RT untuk konsumsi susu

Hasil Uji Beda Karakteristik Wilayah

- Variabel Prediktor Kontinu (Uji t)

Variabel	Statistik uji t	P-value	Keterangan
Pendapatan rumah tangga (X_1)	-11.959	0.000	Beda
Persentase pengeluaran pangan (X_3)	10.976	0.000	Beda
Jumlah anggota rumah tangga (X_4)	-11.683	0.000	Beda
Persentase ART bekerja (X_5)	8.234	0.000	Beda
Persentase ART usia <12 tahun (X_6)	0.177	0.860	Tidak berbeda
Pengeluaran per kapita (X_7)	-5.714	0.000	Beda
Konsumsi susu (Y)	-10.663	0.000	Beda

- Variabel Prediktor Kategorik (Uji chi-square)

Variabel	Statistik uji chi-square	P-value	Keterangan
Pendidikan menengah-rendah (X_{21})	163.656	0.000	Beda
Pendidikan atas-rendah (X_{22})	73.186	0.000	Beda

Tujuan 3: Pemodelan Regresi TKB bagi Pengeluaran RT untuk konsumsi susu

Hasil Uji Linieritas Variabel Prediktor

- Variabel Prediktor Kontinu (Uji RESET)

Variabel Prediktor	Wilayah	Statistik Uji	P-value	Keterangan
Pendapatan rumah tangga (X_1)	Pedesaan	7.077	0.008	Nonlinier
	Perkotaan	48.131	0.000	Nonlinier
Persentase pengeluaran pangan (X_3)	Pedesaan	0.003	0.958	Linier
	Perkotaan	3.683	0.055	Linier
Jumlah anggota rumah tangga (X_4)	Pedesaan	0.769	0.381	Linier
	Perkotaan	0.000	0.639	Linier
Persentase ART bekerja (X_5)	Pedesaan	0.000	0.983	Linier
	Perkotaan	0.644	0.422	Linier
Persentase ART <12 tahun (X_6)	Pedesaan	3.558	0.059	Linier
	Perkotaan	0.644	0.422	Linier
Pengeluaran per kapita (X_7)	Pedesaan	6.201	0.013	Nonlinier
	Perkotaan	24.397	0.000	Nonlinier

Variabel X_7 Pedesaan dan X_1 perkotaan berhasil dilinierkan melalui transformasi logaritma dengan p-value masing-masing 0.317 dan 0.115

- Variabel Prediktor Kategorik (Uji Korelasi Biserial)

Variabel Prediktor	Wilayah	Korelasi Biserial	Statistik Uji t	P-value	Keterangan
Pendidikan menengah-rendah (X_{21})	Pedesaan	-0.096	-3.824	0.000	Linier
	Perkotaan	-0.022	-0.758	0.224	Nonlinier
Pendidikan tinggi-rendah	Pedesaan	-0.178	-7.186	0.000	Linier
	Perkotaan	-0.101	-6.817	0.000	Linier

Tujuan 3: Pemodelan Regresi TKB bagi Pengeluaran RT untuk konsumsi susu

Estimasi Parameter Model Regresi TKB Pedesaan

Parameter	Estimator TKB				
	Kuantil				
	0.05	0.25	0.50	0.75	0.95
β_0	-0.464	0.8988	-0.741	0.275	0.735
β_{21}	-2.633	0.568	0.400	0.862	3.056
β_{22}	1.288	-4.3081	0.460	0.594	0.775
β_3	-4.103	-2.2482	-0.034	0.001	-0.044
β_4	0.374	-0.3244	0.250	0.554	4.430
β_5	-2.479	-4.5692	-0.164	-0.054	-0.283
β_6	-5.281	-2.7559	-0.068	-0.076	-0.335
β_7	0.082	2.7235	1.860	2.129	8.068

Tujuan 3: Pemodelan Regresi TKB bagi Pengeluaran RT untuk konsumsi susu

Estimasi Parameter Model Regresi TKB Perkotaan

Parameter	Estimator TKB				
	Kuantil				
	0.05	0.25	0.50	0.75	0.95
β_0	-4.229	-0.043	0.624	-0.363	-1.021
β_1	-1.989	2.929	1.663	4.351	12.829
β_{21}	1.012	0.206	0.606	0.545	1.910
β_{22}	3.056	0.941	0.280	1.575	3.385
β_3	-5.158	-0.041	-0.062	-0.138	-0.388
β_4	2.668	0.481	0.101	1.818	4.854
β_5	-4.562	-1.341	-0.056	-0.176	-0.413
β_6	-0.148	-0.068	-0.054	-0.213	-0.343

Tujuan 3: Pemodelan Regresi TKB bagi Pengeluaran RT untuk konsumsi susu

Bayes Faktor Model Regresi TKB Reduksi Pedesaan

Model Reduksi (M_2)	Kuantil				
	0.05	0.25	0.50	0.75	0.95
Pend. Menengah-rendah (X_{21})	0.01 ⁽⁰⁾	0.01 ⁽⁰⁾	136.39 ^(**)	753.67 ^(**)	821.6 ^(**)
Pend. Tinggi-rendah (X_{22})	0.01 ⁽⁰⁾	0.01 ⁽⁰⁾	14.29 ^(**)	192.49 ^(**)	49.21 ^(**)
Persentase Pengeluaran pangan (X_3)	0.01 ⁽⁰⁾	0.01 ⁽⁰⁾	-71.84 ⁽⁻⁾	-26.33 ⁽⁻⁾	-180.24 ⁽⁻⁾
Jumlah anggota rumah tangga (X_4)	0.01 ⁽⁰⁾	0.01 ⁽⁰⁾	222.32 ^(**)	307.48 ^(**)	521.86 ^(**)
Persentase ART bekerja (X_5)	0.01 ⁽⁰⁾	0.01 ⁽⁰⁾	90.42 ^(**)	52.03 ^(**)	144.32 ^(**)
Persentase ART usia <12 tahun (X_6)	0.01 ⁽⁰⁾	0.01 ⁽⁰⁾	28.47 ^(**)	231.29 ^(**)	869.97 ^(**)
Pengeluaran per kapita (X_7)	0.01 ⁽⁰⁾	0.01 ⁽⁰⁾	209.27 ^(**)	830.28 ^(**)	1042.66 ^(**)

Keterangan: Kekuatan pembuktian terhadap M_1 , (-) : negatif; (0) : tidak ada; (+) : positif; (*) : kuat; (**): sangat kuat

$$\hat{y}_{(0.75)} = \begin{cases} -0.634 - 0.520X_{21} + 0.935X_{22} + 0.616X_4 - 0.065X_5 - 0.081X_6 + 2.473\log(X_7) & y^* > 0 \\ 0 & y^* \leq 0 \end{cases}$$

Tujuan 3: Pemodelan Regresi TKB bagi Pengeluaran RT untuk konsumsi susu

Bayes Faktor Model Regresi TKB Reduksi Perkotaan

Model Reduksi (M_2)	Kuantil				
	0.05	0.25	0.50	0.75	0.95
Pendapatan rumah tangga (X_1)	0.005 ⁽⁰⁾	37.74 ^(**)	434.52 ^(**)	885.07 ^(**)	1113.55 ^(**)
Pend. Menengah-rendah (X_{21})	0.005 ⁽⁰⁾	37.74 ^(**)	106.53 ^(**)	604.53 ^(**)	644.52 ^(**)
Pend. Tinggi-rendah (X_{22})	0.005 ⁽⁰⁾	24.9 ^(**)	120.08 ^(**)	188.56 ^(**)	186.76 ^(**)
Persentase Pengeluaran pangan (X_3)	0.005 ⁽⁰⁾	9.44 ^(*)	-44.03 ⁽⁻⁾	28.72 ^(**)	32.36 ^(**)
Jumlah anggota rumah tangga (X_4)	0.005 ⁽⁰⁾	20.49 ^(**)	-22.08 ⁽⁻⁾	-36.85 ⁽⁻⁾	-56.01 ⁽⁻⁾
Persentase ART bekerja (X_5)	0.005 ⁽⁰⁾	28.49 ^(**)	-56.52 ⁽⁻⁾	86.29 ^(**)	45.21 ^(**)
Persentase ART usia <12 tahun (X_6)	0.005 ⁽⁰⁾	37.74 ^(**)	27.72 ^(**)	165.65 ^(**)	86.57 ^(**)

Keterangan: Kekuatan pembuktian terhadap M_1 , (-) : negatif; (0) : tidak ada; (+) : positif; (*) : kuat; (**): sangat kuat

$$\hat{y}_{(0.75)} = \begin{cases} 0.422 + 5.597 \log(X_1) + 0.477 X_{21} + 1.420 X_{22} - 0.118 X_3 - 0.198 X_5 - 0.242 X_6 & y^* > 0 \\ 0 & y^* \leq 0 \end{cases}$$

Kesimpulan dan Saran [1]

Kesimpulan

1. Fungsi likelihood pada persamaan berikut digunakan untuk memperoleh estimator model regresi TKB dengan pendekatan bayesian (teknik MCMC Metropolis-Hastings).

$$L(\mathbf{y}|\boldsymbol{\beta}(\theta)) = \theta^n (1-\theta)^n \exp \left\{ -\rho_{\theta} \sum_{i=1}^n (y_i - \max(\mathbf{x}_i' \boldsymbol{\beta}, 0)) \right\}$$

Di sisi lain, pemilihan model terbaik regresi TKB menggunakan formulasi Bayes faktor sebagai berikut.

$$B_{12} = \exp \left\{ \log \hat{m}(\mathbf{y}|M_1) - \log \hat{m}(\mathbf{y}|M_2) \right\}$$

2. Berdasarkan hasil simulasi, performa estimator regresi TKB lebih baik daripada estimator Powell ketika jumlah sampel yang digunakan lebih dari 1000 dan prediktor yang digunakan cukup banyak. Di sisi lain, kedua estimator tersebut ternyata tidak cukup baik untuk memodelkan kuantil bawah seperti kuantil 0.05. Selain itu, performa estimator TKB dan Powell lebih baik daripada estimator tobit standar untuk model yang memiliki error berdistribusi bukan normal.

Kesimpulan dan Saran [2]

Kesimpulan

3. Model regresi TKB terbaik untuk pengeluaran konsumsi susu di wilayah pedesaan yaitu (1) model penuh untuk kuantil 0.05 dan 0.25; (2) model tanpa prediktor persentase pengeluaran pangan untuk kuantil 0.50, 0.75, dan 0.95. Adapun model terbaik bagi wilayah perkotaan yakni (1) model penuh untuk kuantil 0.05 dan 0.25; (2) model tanpa prediktor persentase pengeluaran pangan, jumlah anggota rumah tangga, dan persentase ART bekerja untuk kuantil 0.50; (3) model tanpa prediktor jumlah anggota rumah tangga untuk kuantil 0.75 dan 0.95.

Kesimpulan dan Saran [3]

Saran

1. Pada penelitian ini, model tobit yang digunakan bersifat linier sehingga prediktor yang bersifat nonlinier namun memiliki peranan penting pendapatan rumah tangga tidak dapat dimasukkan dalam model. Oleh karena itu diperlukan pengembangan terhadap model regresi TKB nonlinier
2. Dari hasil simulasi dan analisis yang telah dilakukan, ditemukan bahwa estimator regresi TKB tidak cukup baik untuk mengestimasi model untuk kuantil-kuantil bawah. Oleh karena itu disarankan untuk melakukan modifikasi dan kajian lebih lanjut terkait hal tersebut.

Daftar Pustaka [1]

- Agresti, A. (2002). *Categorical Data Analysis*. New York: John Wiley & Sons.
- Aisyah, N., Arumugam, N., Hussein, M. A., dan Latiff, I. (2012). "Factors Affecting the Technical Efficiency Level of Inshore Fisheries in Kuala Trengganu Malaysia". *International Journal of Agriculture Management and Development*, Vol. 1, 49-56.
- Alwis, A. E. N. D., Ediringhe, J. C., dan Athauda, A. M. T. P. (2009). "Analysis of Factors Affecting Fresh Milk Consumption Among The Mid-Country Consumers". *Tropical Agricultural Research Extension*, Vol. 12, 101-107.
- Babolian, H. R., dan Karim, M. S. A. (2010). "Factors Affecting Milk Consumption Among School Children In Urban and Rural Areas of Selangor Malaysia". *International Food Research Journal*, Vol. 17, 591-601.
- Barnes, R., dan Gillingham, R. (2010). "Demographic Effects in Demand Analysis: Estimation of The Quadratic Expenditure System Using Microdata". *The Review of Economics and Statistics*, Vol. 66, 591-601.
- Biliyas, Y., Chen, S., dan Ying, Z. (2000). "Simple Resampling Methods for Censored Regression Quantiles". *Journal of Econometrics*, Vol. 68, 303-338.
- Buchinsky, M. (1994). "Changes in U.S. Wage Structure 1963-1987: Applications of Quantile Regression". *Econometrica*, Vol. 62, 405-458.
- Buhai, I. S. (2005). "Quantile Regression: Overview and Selected Applications". *Ad Astra*, Vol. 4, 1-17.
- Casella, G., dan Berger, R. L. (2002). *Statistical Inference Second Edition*. California: Duxbury Press.
- Chai, T., dan Draxler, R. R. (2014). "Root Mean Square Error (RMSE) or Mean Absolute Error (MAE): Arguments Against Avoiding RMSE in The Literature". *Geoscientific Model Development*, Vol. 7, 1247-1250.
- Chen, C. (2005). An Introduction to Quantile Regression and The QUANTREG Procedure. from <http://www2.sas.com/proceedingd/sugi30/213-30.pdf>

Daftar Pustaka [2]

- Chib, S., dan Jeliakoz, I. (2001). "Marginal Likelihood from Metropolis-Hastings Algorithm". *Journal of The American Statistical Association*, Vol. 96, 270-281.
- Fitzenberger, B. (1997). *Computational Aspects of Censored Quantile Regression*. Paper presented at the 3rd International Data Analysis based on L- 1-Statistical Procedures and Related Methods, Hayward, California.
- Fitzenberger, B., dan Winker, P. (2007). "Improving the Computation of Censored Quantile Regression Estimators". *CSDA*, Vol. 52, 88-108.
- Gelman, A., Carlin, J. B., Stern, H. S., dan Rubin, D. B. (2004). *Bayesian Data Analysis Second Edition*. Florida: Chapman and Hall.
- Gheblawi, M., dan Sherif, S. (2007). "Determination of Factors Affecting Expenditures on Three Major Food Groups in Al-Ain The United Arab Emirates (UAE)". *Emirates Journal of Food and Agriculture*, Vol. 19, 15-23.
- Greene, W. H. (2001). *Econometrics Analysis 5th Edition*. New Jersey: Prentice Hall.
- Hosmer, D. W., dan Lemeshow, S. (2000). *Applied Logistic Regression*. New York: John Wiley & Sons.
- Jeffreys, H. (1961). *Theory of Probability Third Edition*. Oxford: Oxford Univ. Press.
- Kass, R. E., dan Raftery, A. E. (1995). "Bayes Factors". *Journal of The American Statistical Association*, Vol. 90, 773-795.
- Koenker, R., dan Basset, G. (1978). "Regression Quantiles". *Econometrica*, Vol. 46, 33-50.
- Koenker, R., dan Hallock, K. (2001). "Quantile Regressions". *Journal of Economic Perspectives*, Vol. 15, 143-156.
- Koenker, R., dan Machado, J. A. F. (1999). "Goodness of Fit and Related Inference Processes for Quantile Regression". *Journal of The American Statistical Association*, Vol. 94, 1296-1310.
- Komrattanapanya, P. (2013). "Factors Influencing Dividen Payment in Thailand: A Tobit Regression Analysis". *International Journal of Accounting and Financial Reporting*, Vol. 3, 225-268.

Daftar Pustaka [3]

- Kusfiva, E. (2000). *Analisis Regresi Linier Berganda Dengan Metode OLS, Probit dan Tobit*. (Skripsi), Institut Teknologi Sepuluh Nopember Surabaya, Surabaya.
- Lavine, M., dan Schervish, M. J. (1999). "Bayes Factors: What They Are and What They Are Not". *The American Statistician*, Vol. 53, 119-122.
- Lehmann, E. L., dan Casella, G. (1999). *Theory of Point Estimation*. New York: Springer-Verlag Inc.
- Leiker, A. (2012). *A Comparison Study On The Estimation In Tobit Regression Model* (D. o. S. C. A. a. Science, Trans.). Kansas: Kansas State University.
- McBee, M. (2010). "Modelling Outcomes With Floor or Ceiling Effects: An Introduction to Tobit Model". *Gifted Child Quarterly*, Vol. 54, 313-320.
- Phuong, N. V., Mergenthaler, M., dan Tran, C. H. (2013). *Effects of Socio-economic and Demography Variables on Vietnamese Household' Expenditure for Dairy Product*. Paper presented at the The Tropentag Symposium: AGricultural Development Within The Rural-Urban Continuum Stuttgart.
- Powell, J. (1986). "Censored Regression Quantiles". *Journal of Econometrics*, Vol. 32, 143-155.
- Rahardja, P., dan Manurung, M. (2008). *Pengantar Ilmu Ekonomi (Mikroekonomi & Makroekonomi Edisi Ketiga)*. Jakarta: FE UI.
- Sekhampu, T. J. (2012). "Socio-economic Determinants of Household Food Expenditure in A Low Income Township in South Africa". *Mediterranean Journal of Social Sciences*, Vol. 3, 449-453.
- Sukirno, S. (2004). *Teori dan Pengantar Makro Ekonomi*. Jakarta: Grafindo Persada.
- Tobin, J. (1958). "Estimation of Relationship for Limited Dependent Variables". *Econometrica*, Vol. 26, 24-36.
- Trung, T. Q., Giam, D. Q., Hai, V. T., Thao, L. P., Hang, N. T. T., Son, L. T. K., dan Linh, B. T. M. (2014). "Factors Influencing Milk Consumption of Rural Households in Northern Vietnam". *Greener Journal of Business and Management Studies*, Vol. 4, 31-40.
- USDA. (2013). Indonesian Dairy and Product Annual 2013. from http://gain.fas.usda.gov/Recent%20GAIN%20Publications/Dairy%20and%20Products%20Annual_Jakarta_Indonesia_11-22-2013.pdf

Daftar Pustaka [4]

- Uzunoz, M., dan Akcay, Y. (2012). "A Case Study of Probit Model Analysis of Factors Affecting Consumption of Packed and Unpacked Milk in Turkey". *Economic Research International Journal*, Vol. 1, 1-8.
- Wang, M., dan Zhang, L. (2012). "A Bayesian Quantile Regression Analysis of Potential Risk Factors for Violent Crimes in USA". *Open Journal of Statistics*, Vol. 2, 73-78.
- Wham, C. A., dan Worsley, A. (2003). "New Zealanders' Attitudes To Milk: Implication for Public Health". *Public Health Nutritions*, Vol. 6, 73-78.
- Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Massachusetts: MIT Press.
- Yao, Y. (2010). *Simplex Techniques for Quantile Regression Model Selection*.
- Yu, K., dan Moyeed, R. A. (2001). "Bayesian Quantile Regression". *Statistics and Probability Letters*, Vol. 54, 437-447.
- Yu, K., dan Stander, J. (2007). "Bayesian Analysis of Tobit Quantile Regression Model". *Journal of Econometrics*, Vol. 137, 260-276.
- Yue, Y. R., dan Hong, H. G. (2012). "Bayesian Tobit Quantile Regression Model for Medical Expenditure Panel SURvey Data". *Statistical Modelling*, Vol. 12, 323-346.
- Zain, I. (1997). *Model Regresi Tobit dan Aplikasinya*. Surabaya: Lembaga Penelitian Institut Teknologi Sepuluh Nopember

The background of the slide features a repeating pattern of lotus flowers and gears. The lotus flowers are yellow with blue outlines, and the gears are light blue. The pattern is arranged in a grid-like fashion, with the lotus flowers centered within the gear outlines.

TERIMA KASIH

LAMPIRAN 1

- Model Regresi TKB Terbaik Pedesaan

$$\hat{y}_{(0.25)} = \begin{cases} -4.232 - 0.499X_{21} + 0.659X_{22} - 4.889X_3 - 1.422X_4 - 2.590X_5 - 0.879\log(X_7) & y^* > 0 \\ 0 & y^* = 0 \end{cases}$$

$$\hat{y}_{(0.50)} = \begin{cases} -12.893 + 3.347X_{21} + 8.133X_{22} + 0.404X_4 + 2.307\log(X_7) & y^* > 0 \\ 0 & y^* = 0 \end{cases}$$

$$\hat{y}_{(0.75)} = \begin{cases} -20.869 + 2.646X_{21} + 9.541X_{22} + 1.023X_4 + 5.121\log(X_7) & y^* > 0 \\ 0 & y^* = 0 \end{cases}$$

$$\hat{y}_{(0.95)} = \begin{cases} 0.136 + 3.072X_{21} + 3.905X_{22} - 0.059X_3 - 0.387X_5 + 12.230\log(X_7) & y^* > 0 \\ 0 & y^* = 0 \end{cases}$$

LAMPIRAN 2

- Model Regresi TKB Terbaik Perkotaan

$$\hat{y}_{(0.25)} = \begin{cases} 0.136 + 4.609X_{21} - 0.593X_{22} - 0.821X_3 + 1.714X_4 - 2.834X_5 & y^* > 0 \\ 0 & y^* = 0 \end{cases}$$

$$\hat{y}_{(0.50)} = \begin{cases} -1.597 + 4.743X_{21} + 17.849X_{22} + 0.931X_4 & y^* > 0 \\ 0 & y^* = 0 \end{cases}$$

$$\hat{y}_{(0.75)} = \begin{cases} -3.013 + 8.114X_{21} + 37.095X_{22} + 3.471X_4 & y^* > 0 \\ 0 & y^* = 0 \end{cases}$$

$$\hat{y}_{(0.95)} = \begin{cases} 14.598 + 13.155X_{21} + 50.042X_{22} + 9.147X_4 & y^* > 0 \\ 0 & y^* = 0 \end{cases}$$